

EXCERPT OF REFERENCE 1

Japanese Patent Laid-open No. H11-328194

Laid-open on November 30, 1999

Japanese Patent Application No. H10-130351

Filed on May 31, 1999

Title of the Invention: Keyword Search Method and Apparatus

Description of the Invention:

The present invention relates to a keyword search method and apparatus for converting a user inputted search condition to a search condition more suitable to a user's intention.

Fig. 3 shows a flowchart for explaining the present invention. In step S110, keywords contained in a search condition inputted by a user using a WWW browser are converted to keywords more suitable to a search. In step S120, the converted keywords are presented to the user to urge the user to select a keyword which the user considers to be more appropriate. In step S130, a new search condition is created using the keywords selected by the user in step S120. In step S140, a search is conducted using the created search condition.

Fig. 4 shows a screen generated on a WWW browser comprising regions 310-350. Input Region 310 is for inputting a new search condition by a user. Region 320 is for displaying a search condition used in a previous search. Region 330 is for displaying keywords derived from the keywords contained in the previous search condition. The user can check the boxes to select keywords to be used in a current search. Region 340 is for indicate whether the current search is a re-search or not. Region 350 is for displaying a result of the previous search.

Reference 1

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-328194

(43) 公開日 平成11年(1999)11月30日

(51) Int.Cl.⁸

識別記号

F I

G 0 6 F 17/30

G 0 6 F 15/403

3 2 0 C

3 4 0 C

審査請求 未請求 請求項の数16 O L (全 12 頁)

(21) 出願番号 特願平10-130351

(22) 出願日 平成10年(1998) 5月13日

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 竹野 浩

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(72) 発明者 富田 肇二

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

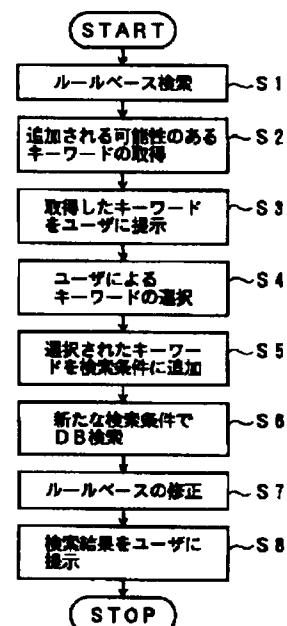
(54) 【発明の名称】 キーワード検索方法及び装置及びキーワード検索プログラムを格納した記憶媒体

(57) 【要約】 (修正有)

【課題】 ユーザが入力した検索条件をユーザの意図に、より適していると思われる検索条件に変換し、キーワード検索を支援する。

【解決手段】 検索で使用する任意のキーワードと、任意のキーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベースを検索して、ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、ユーザが入力した検索条件に含まれないキーワードを取得しユーザに表示して、キーワードから実際の検索に使用するキーワードを選択させ、ユーザが入力した検索条件に追加して、検索を実行する。

本発明の原理を説明するための図



【特許請求の範囲】

【請求項1】 ユーザが入力したキーワードまたは、キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索方法において、

前記検索で使用する任意のキーワードと、該任意のキーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベースを検索して、前記ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、該ルールベースに登録される可能性のあるキーワードを取得し、

前記ルールベースより取得したキーワードを前記ユーザに表示し、該キーワードから実際の検索に使用するキーワードを選択させ、

前記ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加して、検索を実行し、

前記ユーザが選択したキーワードに基づいて前記ルールベースを修正し、

前記検索の結果を前記ユーザに表示することを特徴とするキーワード検索方法。

【請求項2】 前記ユーザが入力したキーワードが複数である場合に、

前記キーワードの関連度に関する情報が前記ルールベースに存在しない組み合わせに関しては、予め定められた値を関連度として、該ルールベースに保存し、存在する組み合わせに関しては、該ルールベースに対してこれらのキーワードの関連度を高める修正処理を行う請求項1記載のキーワード検索方法。

【請求項3】 前記ユーザが入力した前記キーワードと該ユーザが選択したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が前記ルールベースに存在する組み合わせに関しては、該ルールベースに対して、該キーワード間の関連度を高める修正処理を行う請求項1記載のキーワード検索方法。

【請求項4】 前記ユーザが入力した前記キーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、前記ルールベースに存在する組み合わせに関しては、該ルールベースに対して該キーワード間の関連度を低くする修正処理を行う請求項1記載のキーワード検索方法。

【請求項5】 前記修正処理が行われたルールベースを用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行う請求項1乃至4記載のキーワード検索方法。

【請求項6】 ユーザが入力したキーワードまたは、キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索装置であって、

前記検索で使用する任意のキーワードと、該任意のキ

ーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベースと、

ユーザからの検索条件の人力受け付けるユーザ検索条件入力手段と、

検索の対象となる検索対象データベースと、

前記ユーザ検索条件入力手段により入力された前記検索条件に基づいて前記検索対象データベースを検索する検索手段と、

10 前記ルールベースの検索を行い、前記ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、該ルールベースに登録される可能性のあるキーワードを取得する他キーワード取得手段と、

前記ルールベースより取得したキーワードをユーザに表示し、該キーワードから実際の検索に使用するキーワードを選択させるユーザ選択指示手段と、

前記ユーザ選択指示手段により、前記ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加し

20 て、検索を実行する再検索手段と、
前記選択キーワードに基づいて前記ルールベースを修正するルールベース修正手段と、

前記検索の結果を前記ユーザに表示する検索結果表示手段とを有することを特徴とするキーワード検索装置。

【請求項7】 前記ルールベースは、
入力としての入力キーワード、該入力キーワードから派生する派生キーワード及び該入力キーワードと該派生キーワード間の関連度から構成される請求項6記載のキーワード検索装置。

30 【請求項8】 前記ルールベース修正手段は、
前記ユーザが入力したキーワードが複数である場合に、前記キーワード間の関連度に関する情報が前記ルールベースに存在しない組み合わせに関しては、予め定められた値を関連度として、該ルールベースに保存し、存在する組み合わせに関しては、該ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う第1のルールベース修正手段を更に有する請求項6記載のキーワード検索装置。

40 【請求項9】 前記ルールベース修正手段は、
前記ユーザが入力した前記キーワードと該ユーザが選択したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が前記ルールベースに存在する組み合わせに関しては、該ルールベースに対して、該キーワード間の関連度を高める修正処理を行う第2のルールベース修正手段を更に有する請求項6記載のキーワード検索装置。

50 【請求項10】 前記ルールベース修正手段は、
前記ユーザが入力した前記キーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、前記ルールベ...

スに存在する組み合わせに関しては、該ルールベースに対して該キーワード間の関連度を低くする修正処理を行う第3のルールベース修正手段を更に有する請求項6記載のキーワード検索装置。

【請求項11】 前記検索手段は、

前記修正処理が行われたルールベースを用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行う手段を含む請求項8乃至10記載のキーワード検索装置。

【請求項12】 ユーザが入力したキーワードまたは、
10 キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索プログラムを格納した記憶媒体であって、
ユーザからの検索条件の入力受け付けるユーザ検索条件入力プロセスと、

前記ユーザ検索条件入力プロセスにより入力された前記検索条件に基づいて前記検索対象データベースを検索する検索プロセスと、

前記検索プロセスで使用される任意のキーワードと、該任意のキーワードを含む任意のキーワードとの関連の度
20 合いを示す情報を、検索、保存、修正することが可能なルールベースを検索するルールベース検索プロセスと、
前記ルールベース検索プロセスにより検索を行い、前記ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、該ルールベースに登録される可能性のあるキーワードを取得する他キーワード取得プロセスと、

前記ルールベースより取得したキーワードをユーザに表示し、該キーワードから実際の検索に使用するキーワードを選択させるユーザ選択指示プロセスと、
30 前記ユーザ選択指示プロセスにより、前記ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加して、検索対象データベースの検索を実行する再検索プロセスと、

前記選択キーワードに基づいて前記ルールベースを修正するルールベース修正プロセスと、
前記検索の結果を前記ユーザに表示させる検索結果表示プロセスとを有することを特徴とするキーワード検索プログラムを格納した記憶媒体。

【請求項13】 前記ルールベース修正プロセスは、
40 前記ユーザが入力したキーワードが複数である場合に、前記キーワード間の関連度に関する情報が前記ルールベースに存在しない組み合わせに関しては、予め定められた値を関連度とて、該ルールベースに保存し、存在する組み合わせに関しては、該ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う第1のルールベース修正プロセスを更に有する請求項12記載のキーワード検索プログラムを格納した記憶媒体。

【請求項14】 前記ルールベース修正プロセスは、
50 前記ユーザが入力した前記キーワードと該ユーザが選択

したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が前記ルールベースに存在する組み合わせに関しては、該ルールベースに対して、該キーワード間の関連度を高める修正処理を行う第2のルールベース修正プロセスを更に有する請求項12記載のキーワード検索プログラムを格納した記憶媒体。

【請求項15】 前記ルールベース修正プロセスは、前記ユーザが入力した前記キーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、前記ルールベースに存在する組み合わせに関しては、該ルールベースに対して該キーワード間の関連度を低くする修正処理を行う第3のルールベース修正プロセスを更に有する請求項12記載のキーワード検索プログラムを格納した記憶媒体。

【請求項16】 前記検索プロセスは、前記修正処理が行われたルールベースを用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行うプロセスを含む請求項12乃至15記載のキーワード検索プログラムを格納した記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、キーワード検索方法及び装置及びキーワード検索プログラムを格納した記憶媒体に係り、特に、キーワード検索作業において、ユーザが入力した検索条件をより適切と思われるものに変換することによって、情報検索作業の効率を高めるためのキーワード検索方法及び装置及びキーワード検索プログラムを格納した記憶媒体に関する。

【0002】

【従来の技術】従来のキーワード検索方法として、以下のような方法がある。

1. ユーザが入力した検索条件をそのまま用いてキーワード検索を実施する。
2. 単語と単語の関係を記述したシソーラスと呼ばれる辞書を予め、手動で、または同一の文書集合に現れる単語同士は関連があるとする共起関係を用いて検索対象の文書から自動的に作成する。以後、検索条件として、あるキーワードが入力されると、シソーラス内にある関連語を静的に追加することによって、最初の検索条件をより適切と思われるものに変換してキーワード検索を実施する。

【0003】3. ユーザの検索作業を監視して、複数のユーザがそれぞれ異なったキーワードを用いて検索を行い、ある1つの情報を正しい検索結果として選択した場合、それぞれの検索に使われたキーワードが互いに関連があると判断し、これらを関連語として保存する。以後、検索条件として、これらのキーワードのどれかが入力されると、関連語を追加することによって、最初の検索条件をより適切と思われるものに変換してキーワード

検索を実施する。

【0004】

【発明が解決しようとする課題】しかしながら、上記従来のキーワード検索方法には以下のような問題が生じる。

1. ユーザが入力した検索条件をそのまま用いてキーワード検索を実施する方式では、ユーザが入力したキーワードと検索対象が持つキーワードの表記に差がある場合（例：コンピューターとコンピュータ）や、同義語である場合（例：コンピュータと計算機）に、適切な検索が

できないという問題が生じる。
【0005】2. 手動で作成したシソーラスを用いて、入力された検索条件中のキーワードと関係の深い単語を静的に派生させ、これを用いてこの検索条件を変換して（例：コンピュータを、“コンピューターV計算機”に変換）キーワード検索を実施する方式では、上記の表記のゆれ、同義語の問題をある程度解決することができる。また、共起関係を用いてシソーラスを自動生成し、これを用いて検索条件を変換する方式では、ユーザにとって有用な検索条件の変換を、ある程度行うことができる。例えば、“コンピューター”と“グラフィックス”を共に含む文書が検索対象の文書中に数多くあったならば、“コンピューター”と“グラフィックス”の間に共起関係が成立するので、この関係をシソーラスに登録することにより、“コンピューター”を“コンピューターハグラフィックス”に変換してキーワード検索を実施することができる。しかしながら、シソーラスにすべての単語同士の関係を手動で記述することは到底不可能である。また、検索対象の文書内に共起関係が成り立たないものについては、関連語を派生させる規則を自動生成することができない。例えば、あるユーザがCPUの交換に関する文書を探したい場合でも検索対象の文書の中に“CPU”と“交換”を共に含む文書が多数なければ、“CPU”から“交換”を派生させる規則は生成されない。このように予め作成した静的なシソーラスを使用したのでは、適切な関連語の派生を行うことができないという問題が生じる。

【0006】3. 複数のユーザが同じ情報を検索するために入力した異なったキーワードを互いに関連があると判断し、これを用いて検索条件を変換する方式では、変換のための規則が動的に生成されるので、上記の問題をある程度解決することができる。しかしながら、検索の結果、選択された情報が複数の内容を含んでおり、複数のユーザが入力した検索条件がそれぞれ異なった内容に対するものであった場合、使用された検索条件は互いに関連がない可能性が高いにもかかわらず、一律に関連があると判断されて保存されてしまう。そのためこの方式で検索条件を変換した場合、ユーザの意図と異なった検索条件になってしまう危険がある。

【0007】例えば、選択された情報がコンピュータグラフィックスを作成するプログラムの使用説明書であって、このプログラムを起動するためのオペレーションシステムの操作に関する説明と、コンピュータグラフィックスによるデザインの手法に関する説明を含んでいたとする。この場合、この方式では、オペレーションシステムに関するキーワードと、デザインに関するキーワードが互に関連があると判断されてしまう可能性があり、以後、デザインに関する情報の検索を行う際に、オペレーションシステムに関する情報がヒットしてしまうという不都合が生じてしまう。このように検索条件の誤った変換規則が保存されてしまうという問題が生じる。

【0008】上記のように、従来の技術によるキーワード検索には、ユーザが入力したキーワードが適切なものでない場合、適切な検索を実行することができない、ユーザの検索を支援する機能が必ずしも適切に働かないという問題がある。本発明は、上記の点に鑑みなされたもので、従来技術によるユーザの入力した検索条件をそのまま用いてキーワード検索を実施するキーワード検索装置とユーザとの間に介在し、ユーザが入力した検索条件をユーザの意図に、より適していると思われるキーワードを含む検索条件に変換し、キーワード検索を支援することが可能なキーワード検索方法及び装置及びキーワード検索プログラムを格納した記憶媒体を提供することを目的とする。

【0009】

【課題を解決するための手段】図1は、本発明の原理を説明するための図である。本発明（請求項1）は、ユーザが入力したキーワードまたは、キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索方法において、検索で利用される任意のキーワードと、該任意のキーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベースを検索して（ステップ1）、ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、ルールベースに登録される可能性があるキーワードを取得し（ステップ2）、ルールベースより取得したキーワードをユーザに表示し（ステップ3）、該キーワードから実際の検索に使用するキーワードを選択させ（ステップ4）、ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加して（ステップ5）、検索を実行し（ステップ6）、ユーザが選択したキーワードに基づいてルールベースを修正し（ステップ7）、検索の結果をユーザに表示する（ステップ8）。

【0010】本発明（請求項2）は、ユーザが入力したキーワードが複数である場合に、キーワード間の関連度に関する情報がルールベースに存在しない組み合わせに関しては、予め定められた値を関連度として、該ルール

ベースに保存し、存在する組み合わせに関しては、該ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う。

【0011】本発明（請求項3）は、ユーザが入力したキーワードと該ユーザが選択したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報がルールベースに存在する組み合わせに関しては、該ルールベースに対して、該キーワード間の関連度を高める修正処理を行う。本発明（請求項4）は、ユーザが入力したキーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、ルールベースに存在する組み合わせに関しては、該ルールベースに対して該キーワード間の関連度を低くする修正処理を行う。

【0012】本発明（請求項5）は、修正処理が行われたルールベースを用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行う。図2は、本発明の原理構成図である。本発明は（請求項6）は、ユーザが入力したキーワードまたは、キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索装置であって、検索で使用される任意のキーワードと、該任意のキーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベース290と、ユーザからの検索条件の入力受け付けるユーザ検索条件入力手段200と、検索の対象となる検索対象データベース280と、ユーザ検索条件入力手段200により入力された検索条件に基づいて検索対象データベース280を検索する検索手段250と、ルールベース290の検索を行い、ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、ルールベース290に登録される可能性があるキーワードを取得する他キーワード取得手段270と、ルールベース290より取得したキーワードをユーザに表示し、該キーワードから実際の検索に使用するキーワードを選択させるユーザ選択指示手段220と、ユーザ選択指示手段220により、ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加して、検索を実行する再検索手段255と、選択キーワードに基づいてルールベース290を修正するルールベース修正手段260と、検索の結果をユーザに表示する検索結果表示手段225とを有する。

【0013】本発明（請求項7）は、ルールベース290が、入力としての入力キーワード、該入力キーワードから派生する派生キーワード及び該入力キーワードと該派生キーワードの関連度から構成される。本発明（請求項8）は、ルールベース修正手段260において、ユーザが入力したキーワードが複数である場合に、キーワード間の関連度に関する情報がルールベース290に存在しない組み合わせに関しては、予め定められた値を関連

度として、該ルールベース290に保存し、存在する組み合わせに関しては、該ルールベース290に対してこれらのキーワード間の関連度を高める修正処理を行う第1のルールベース修正手段を更に有する。

【0014】本発明（請求項9）は、ルールベース修正手段260において、ユーザが入力したキーワードと該ユーザが選択したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報がルールベース290に存在する組み合わせに関しては、該ルールベース290に対して、該キーワード間の関連度を高める修正処理を行う第2のルールベース修正手段を更に有する。

【0015】本発明（請求項10）は、ルールベース修正手段260において、ユーザが入力したキーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、ルールベース290に存在する組み合わせに関しては、該ルールベース290に対して該キーワード間の関連度を低くする修正処理を行う第3のルールベース修正手段を更に有する。

【0016】本発明（請求項11）は、検索手段250において、修正処理が行われたルールベース290を用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行う手段を含む。本発明（請求項12）は、ユーザが入力したキーワードまたは、キーワードをブール演算子で結合した式から構成された検索条件に関する検索を実行するキーワード検索プログラムを格納した記憶媒体であって、ユーザからの検索条件の入力受け付けるユーザ検索条件入力プロセスと、ユーザ検索条件入力プロセスにより入力された前記検索条件に基づいて前記検索対象データベースを検索する検索プロセスと、検索プロセスで使用される任意のキーワードと、該任意のキーワードを含む任意のキーワードとの関連の度合いを示す情報を、検索、保存、修正することが可能なルールベースを検索するルールベース検索プロセスと、ルールベース検索プロセスにより検索を行い、ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつ、ルールベースに登録される可能性があるキーワードを取得する他キーワード取得プロセスと、ルールベースより取得したキーワードをユーザに表示し、該キーワードから実際の検索に使用するキーワードを選択させるユーザ選択指示プロセスと、ユーザ選択指示プロセスにより、ユーザが選択したキーワードを、該ユーザが入力した検索条件に追加して、検索を実行する再検索プロセスと、選択キーワードに基づいてルールベースを修正するルールベース修正プロセスと、検索の結果をユーザに表示させる検索結果表示プロセスとを有する。

【0017】本発明（請求項13）は、ルールベース修正プロセスにおいて、ユーザが入力したキーワードが複

数である場合に、キーワード間の関連度に関する情報がルールベースに存在しない組み合わせに関しては、予め定められた値を関連度とて、該ルールベースに保存し、存在する組み合わせに関しては、該ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う第1のルールベース修正プロセスを更に有する。

【0018】本発明（請求項14）は、ルールベース修正プロセスにおいて、ユーザが入力したキーワードと該ユーザが選択したキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報がルールベースに存在する組み合わせに関しては、該ルールベースに対して、該キーワード間の関連度を高める修正処理を行う第2のルールベース修正プロセスを更に有する。

【0019】本発明（請求項15）は、ルールベース修正プロセスにおいて、ユーザが入力したキーワードと該ユーザが選択しなかったキーワードの全ての組み合わせに対して、該キーワード間の関連度に関する情報が、ルールベースに存在する組み合わせに関しては、該ルールベースに対して該キーワード間の関連度を低くする修正処理を行う第3のルールベース修正プロセスを更に有する。

【0020】本発明（請求項16）は、検索プロセスにおいて、修正処理が行われたルールベースを用いて、次回以降の検索時に、関連したキーワード取得を行い、検索を行うプロセスを含む。上記により、本発明は、従来の技術とは異なり、キーワードを変換する際に、変換規則として、手動または、共起関係を利用して予め作成した静的なシソーラスを用いるのではなく、データベースの検索で使用する任意のキーワードと任意のキーワードの関連の度合いを示す情報を検索、保存、修正する機能をルールベースに持たせ、ユーザが入力した検索条件に含まれるキーワードに対して、関連の度合いが予め定められた値以上であり、かつユーザが入力した検索条件に含まれないキーワードを取得してユーザに提示することが可能である。

【0021】さらに、入力されたキーワードが複数であるときに、これらのキーワード関連度に関する情報がルールベースに存在しない組み合わせに関しては、予め定められた値を関連度としてルールベースに保存し、存在する組み合わせに対しては、ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う。また、ユーザが入力したキーワードとユーザが選択したキーワードの全ての組み合わせに対して、これらのキーワード間の関連度に関する情報が上記のルールベースに存在する組み合わせである場合には、当該ルールベースに対してこれらのキーワード間の関連度を高める修正処理を行う。

【0022】また、ユーザが入力したキーワードとユーザが選択しなかったキーワードの全ての組み合わせに対して、これらのキーワード間の関連度に関する情報がデ

ータベースに存在する組み合わせについては、ルールベースに対してこれらのキーワード間の関連度を低くする修正処理を行う。これにより、以降の検索において、動的な検索条件の変換を行うことが可能となる。

【0023】

【発明の実施の形態】以下の説明において、キーワードをブール演算子で結合した式を「検索条件」、キーワードを検索により適したキーワードに変換する処理を「キーワード派生処理」、キーワード派生処理を行うための変換規則を「キーワード派生ルール」、キーワード派生ルールを動的に修正する処理を「派生ルール修正処理」、派生されたキーワードを用いて検索条件を生成する処理を「検索条件生成処理」と記す。

【0024】最初に、本発明の検索の概要を説明する。図3は、本発明の検索の概要を説明するためのフローチャートである。ステップ110） キーワード派生処理として、ユーザが入力した検索条件に含まれるキーワードを、検索により適したキーワードに変換する。

ステップ120） 派生キーワード選択処理として、キーワード派生処理（ステップ110）で変換されたキーワードを、ユーザに提示し、よりユーザ自身が適当と思うキーワードを選択させる。

【0025】ステップ130） 検索条件生成処理として、派生キーワード選択処理（ステップ120）で、ユーザが選択したキーワードを用いて、新たな検索条件を生成する。

ステップ140） 検索実行処理として、検索条件生成処理（ステップ130）で生成した検索条件を用いて検索を実施する。

【0026】以下の記号を用いて、上記の各処理を詳しく説明する。

U_i ：ユーザの入力した検索条件に含まれるキーワードからなる集合：

R_i ：キーワード検索検索装置に入力する検索条件に含まれるキーワードからなる集合：

A_i ：画面に表示する派生キーワードからなる集合：

S_i ： A_i の中で、ユーザが選択したキーワードからなる集合：

D_i ： A_i の中で、ユーザが選択しなかったキーワードからなる集合：すべての t において、 $A_i = S_i \cup D_i$ 、 $S_i \cap D_i = \phi$ （空集合）である。また、初回の検索（ $t=1$ ）では、 A_i 、 S_i 、 D_i は、 ϕ であり、 $R_i = U_i$ である。

【0027】各処理の前提としてキーワード派生ルールの形式について説明する。キーワード派生ルールは、以下の形式を持つ。

(key1, key2, value)

ここで、key1は、入力としてのキーワード、key2は、key1から派生されるキーワードであり、valueは、key1とkey2の関連度を示す。ある1

11

つのkey1に対して、key2が互いに異なる複数の派生ルールが存在することが可能である。

【0028】最初に、ステップ110のキーワード派生処理について説明する。キーワード派生処理は、 $t-1$ 回目の検索条件に含まれる全てのキーワード R_{t-1} が *

$$\begin{aligned} & (a_{11}, v_{11}) (a_{12}, v_{12}) \cdots (a_{1n}, v_{1n}) \\ & (a_{21}, v_{21}) (a_{22}, v_{22}) \cdots (a_{2n}, v_{2n}) \end{aligned}$$

$$(a_{q1}, v_{q1}) (a_{q2}, v_{q2}) \cdots (a_{qn}, v_{qn})$$

を生成する。ここでは、 a_{pq} は、 key_p から派生された q 番目のキーワードで、 v_{pq} は、 key_p と a_{pq} の関連度である。また、 $(n^{('')})$ は、 key_p が持つ派生ルールの個数である。

【0029】次に、このリストを走査し、 $a_{i1} = a_{pq}$ となる要素が存在すれば、 (a_{i1}, v_{i1}) を $(a_{i1}, v_{i1} + v_{pq})$ として、 (a_{pq}, v_{pq}) をリストから削除する。この方法で生成したリストから v_{i1} の値の高い順に、予め定められた整数 k 個だけ要素を取り出して、キーワードを抽出することにより、 A_t を生成する。次に、前述のステップ120の派生キーワード選択処理について説明する。

【0030】本処理は、キーワード派生処理(ステップ110)で生成したキーワードのリスト A_t をユーザに※

・絞込検索:

$$(t=1) \quad R'_t = U'_t \quad (1)$$

$$(t \geq 2) \quad R'_t = R_{t-1} \wedge (U'_t \vee S'_t) \quad (2)$$

・再検索:

$$(t=1) \quad R'_t = U'_t \quad (3)$$

$$(t \geq 2) \quad R'_t = R_{t-1} \vee (U'_t \vee S'_t) \quad (4)$$

R'_t は、 t 回目の検索において、キーワード検索装置に入力する検索条件である。 R'_{t-1} は、前回の($t-1$ 回目)の検索で用いた検索条件である。 U'_t は、 U_t の要素である"をブール演算子で結合した式である。 S'_t は、 S_t に含まれるすべてのキーワードを \vee で結合した式である。

【0033】次に、ステップ140のキーワード派生ルールの修正処理について説明する。最初に使用する用語を説明する。

・ルールの追加($k_1, k_2, \text{def_value}, \text{value}$): ルールベース内に(k_1, k_2, v)という項目があったならば、この項目を($k_1, k_2, v + \text{value}$)に修正する。

【0034】・ルールの修正(k_1, k_2, value): ルールベース内に(k_1, k_2, v)という項目があったならば、この項目を($k_1, k_2, v + \text{value}$)に修正する。もし、(k_1, k_2, v)という項目がなければ何も行わない。

・ $\text{num}(G)$ 集合(G)に含まれる要素の数。

【0035】システムは、キーワード派生ルールの修正

12

※ら、 A_t を生成する処理である。 R_{t-1} の個々の要素であるキーワード $key_1, key_2, \dots, key_n$ のそれぞれに対して、対応するキーワード派生ルールを摘要し、派生すべきキーワードとその関連度の対のリスト

※提示し、ユーザの検索の目的に適合したキーワードをユーザに選択させる処理である。これにより、 S_t を生成する。また、 A_t と S_t から D_t を生成する。次に、前述のステップ130の検索条件生成処理について説明する。

【0031】 t 回目の検索には、 $t-1$ 回目の検索でヒットした件数が多過ぎるため、検索条件がより厳しくなるように、 $t-1$ 回目の検索条件に検索条件を追加する「絞込検索」と、 $t-1$ 回目の検索ではヒットした件数が少な過ぎるため、検索条件がよりゆるやかになるように、 $t-1$ 回目の検索条件に検索条件を追加する「再検索」の2種類の検索が存在する。

【0032】本発明は、両者を区別し、そのそれぞれに対して検索生成処理を用意する。

を、 R_{t-1}, U_t, S_t, D_t を用いて行う。

1. U_t の中の要素の任意の2個のすべての組み合わせについて、ルールの追加($r_i, u_i, \text{def_value1}, \text{value1}$)を行う。 $(\text{num}(U_t)C_2)$ 個のルールの追加)

2. R_{t-1} と U_t の中の要素のすべての組み合わせについて、ルールの追加($r_i, u_i, \text{def_value2}, \text{value2}$)を行う。

【0036】 $(\text{num}(R_{t-1})C_1 \times \text{num}(U_t)C_1)$ 個のルールの追加)

3. R_{t-1} と S_t の中の要素のすべての組み合わせについて、ルールの修正($r_i, s_i, +\text{value3}$)を行う。 $(\text{num}(R_{t-1})C_1 \times \text{num}(S_t)C_1)$ 個のルールの修正)

4. R_{t-1} と D_t の中の要素のすべての組み合わせについて、ルールの修正($r_i, d_i, -\text{value4}$)を行う。

【0037】 $(\text{num}(R_{t-1})C_1 \times \text{num}(D_t)C_1)$ 個のルールの修正)

最後にステップ150の検索実行処理を説明する。検索

条件生成処理(ステップ130)で生成した検索条件R'を用いて検索対象のデータベースの検索を実行する。なお、当該検索処理は、通常のキーワード検索装置をそのまま使用することが可能である。

【0038】

【実施例】図4は、本発明の一実施例の情報検索装置の構成を示す。以下の例では、インターネットで広く用いられている“WWW”を用いて、情報検索装置を実現しているものである。ユーザは、WWWのブラウザを用いて、検索条件を入力し、入力を実現しているものである。ユーザは、WWWのブラウザを用いて検索条件を入力し、入力した検索条件は、WWWの通信手段を用いて、本発明の情報検索装置に送信される。送信された検索条件は、以下に説明する手順に従って処理される。処理結果は、WWWの通信手段を用いてユーザのWWWブラウザに送信され、ユーザに対して表示されるものである。

【0039】同図に示す情報検索装置は、形態素解析部210、WWW処理部220、検索条件生成部230、情報検索制御部240、検索処理部/結果変換部250、派生ルール修正処理部260、キーワード派生処理部270、検索対象データベース280及びルールベース290から構成される。形態素解析部210は、検索条件生成部230から追加検索条件bを受信し、これを解析し、単語の集合に変換してキーワードを抽出し、その結果を、適切なブール演算子を用いて結合し、検索条件の解析結果dとして、検索条件生成部230に送信する部分である。

【0040】WWW処理部220は、ユーザのWWWブラウザに図5に示す形式で表示を行い、ユーザに操作させることで、前回の検索条件a、追加検索条件b、キーワード選択情報cを受信し、情報検索制御部240に送信し、その結果として、検索結果f、派生キーワードg、検索条件eを受信して、ユーザのWWWブラウザに送信し、表示させる部分である。

【0041】検索条件生成部230は、情報検索制御部240から、前回の検索条件a、追加の検索条件b、キーワード選択情報cを受信し、形態素解析部210を用いて、実際に検索を行う検索条件eを生成し、情報検索制御部240に送信する部分である。情報検索制御部240は、WWW処理部220、検索条件生成部230、検索処理部/結果変換部250、派生ルール修正処理部260、キーワード派生処理部270と交信し、検索条件eの生成、検索の実行、キーワード派生の処理を行う部分である。

【0042】検索処理部/結果変換部250は、情報検索制御部240から、検索条件eを受信して、検索対象データベース280に送信して、検索を実行し、その結果をWWWのブラウザに表示可能な形式に変換したものを、情報検索制御部240に送信する部分である。派生

ルール修正処理部260は、情報検索制御部240から、前回の検索条件a、追加検索条件b、キーワード選択情報cを受信し、キーワード派生ルールを修正し、ルールベース290に保存する部分である。

【0043】キーワード派生処理部270は、情報検索制御部240から、検索条件eを受信し、キーワードの派生処理を実行し、派生されたキーワードgを情報検索制御部240に送信する部分である。検索対象データベース280は、検索処理部/結果変換部250から検索条件を受信し、検索を実行し、検索結果を検索処理部/結果変換部250に送信する部分である。

【0044】ルールベース290は、キーワード派生ルールの保存、取り出し、修正を行うためのデータベースである。キーワード派生ルールは、

(key1, key2, value)

の形式を持つ。key1は、入力としてのキーワード、key2は、key1から派生されるキーワードであり、valueは、key1とkey2の関連度を示す。ルールベース290は、key1を入力することで、key1に関する全ての派生ルールを出力すること、key1, key2を入力することで、これらの単語の関連度valueを出力することも可能である。

【0045】次に、各ブロックでやり取りされる情報の説明を行う。

a：前回の検索条件は、直前の検索で使用された検索条件であり、キーワードとなる単語をブール演算子で結合したものである。1回目の検索では、空文字列が与えられる。

b：追加検索条件は、今回の検索でユーザが新たに入力した検索条件である。自然文の形式に、これが、「絞込検索」であるか、「再検索」であるかの情報を付与したものである。

【0046】c：キーワード選択情報は、ユーザのWWWブラウザに表示された派生キーワード及び、当該キーワードのそれぞれがユーザによって選択されたか否かの情報の集合である。

d：検索条件の解析結果は、ユーザが自然文の形式で入力した追加検索条件bに、形態素解析を行って、キーワードとなる単語を抽出し、ブール演算子で結合したものである。

【0047】e：検索条件は、実際に検索を実行するための検索条件で、キーワードとなる単語をブール演算子で結合したものである。

f：検索結果は、検索を行った結果である。ユーザのWWWブラウザで表示可能な形式で表現される。

g：派生されたキーワードは、ルールベース290内のキーワード派生ルールを用いて派生されたキーワードの集合である。

【0048】図5は、本発明の一実施例の情報検索装置のユーザインタフェースの画面の例を示す。これは、図

4のWWW処理部220が生成するもので、ユーザのWWWブラウザ上に表示されるものである。追加検索条件入力部310は、ユーザが新たな検索条件を自然文の形式で入力する部分である。

【0049】前回検索条件表示部320は、前回の検索に用いられた検索条件の表示を行う部分である。派生キーワード表示・選択部330は、前回の検索条件に含まれるキーワードから派生されたキーワードを表示する部分であり、かつユーザはこれらのキーワードに付与されたチェックボックスをクリックすることにより、今回の検索においてこれらのキーワードを選択する／選択しないを指示することができる。

【0050】検索種別選択部340は、今回の検索が「絞込検索」か「再検索」のどちらであるか指示する部分である。検索結果表示部350は、前回の検索の結果を表示する部分である。これは、WWWリンク情報として表示される。次に、本実施例の動作を順に説明する。

【0051】(1) WWW処理部220は、ユーザより前回の検索条件a、追加検索条件b、キーワード選択情報cを受信し、情報検索制御部240に送信する。

(2) 情報検索制御部240は、受信した情報を検索条件生成部230に送信する。

(3) 検索条件生成部230は、受信した情報の中の追加検索条件bの自然文の形式の情報を形態素解析部210に送信する。

【0052】(4) 形態素解析部210は、受信した自然文を解析し、単語を抽出し、これをキーワードの集合とし、ブール演算子で結合して、検索条件生成部230に送信する。

(5) 検索条件生成部230は、形態素解析部210から受信した検索条件の解析結果dに、上記の(3)で受信した前回の検索条件aと、追加検索条件の中の「絞込検索」、または、「再検索」かの情報、及びキーワード選択情報cの中のユーザが選択したキーワードの集合を用いて、前述の検索条件生成処理(図3のステップ130)を実施し、検索条件eを生成し、情報検索制御部240に送信する。

【0053】(6) 情報検索制御部240は、受信した検索条件eを検索処理部／結果変換部250に送信する。

(7) 検索処理部／結果変換部250は、受信した検索条件eを用いて、検索対象データベース280に対する検索を実行して、検索の結果を受信し、WWWブラウザで表示可能な形式を持つ検索結果fに変換して、情報検索制御部240に送信する。

【0054】(8) 情報検索制御部240は、今回の検索に使用した検索条件eを、キーワード派生処理部270に送信する。

(9) キーワード派生処理部270は、受信した検索条件eに含まれる個々のキーワードに対してルールベ

ス290をアクセスし、前述のキーワード派生処理(図3のステップ110)を実施し、派生されたキーワードgを生成し、情報検索制御部240に送信する。

【0055】(10) 情報検索制御部240は、検索結果fに、今回使用した検索条件e、派生されたキーワードgを加えてWWW処理部220に送信し、図5のユーザインタフェース画面を再度生成し、ユーザのWWWブラウザに表示させる。

(11) 情報検索制御部240は、前回の検索条件a、追加検索条件b、及びキーワード選択情報cを、派生ルール修正処理部260に送信する。派生ルール修正処理部260は、これらを用いて、前述のキーワード派生修正処理(図3のステップ140)を実施し、修正結果をルールベース290内に格納する。

【0056】上述のように、本発明によれば、表記のゆれ、同義語の問題を解決できる。詳しくは、先に例示した、“コンピュータ”、“コンピューター”、“計算機”の例を用いて、本発明が表記のゆれ、同義語の問題を解決する効果を奏することを説明する。

キーワード派生ルールとして、

(コンピューター、コンピュータ、関連度“大”)

(コンピューター、計算機、関連度“大”)

が用意されているものとする。

【0057】ユーザが1回目の検索条件として、“コンピューター”を入力して検索を実行したが、探している情報に含まれているキーワードが“コンピュータ”または、“計算機”であったため、ヒットした情報が0個であったとする。本発明を用いて2回目の検索を行う場合、以下の処理が実行される。

(1) キーワード派生処理(図3、ステップ110)で、“コンピュータ”と“計算機”が生成される。

【0058】(2) 派生キーワード選択処理(図3のステップ120)で、ユーザに対して“コンピュータ”と“計算機”が表示される。ここで、ユーザが双方のキーワードを選択したものとする。

(3) 1回目の検索でヒットした情報が0個であったので、次に行われる検索は、「再検索」である。従って、検索条件生成処理(図3、ステップ130)で、2回目の検索条件として“コンピューターV(コンピュータV計算機)”が生成される。

【0059】(4) 検索実行処理(図3のステップ150)で、“コンピューターV(コンピュータV計算機)”の検索が行われる。これにより、“コンピュータ”または、“計算機”を含む情報を検索結果として得ることができ、表記のゆれ、同義語の問題を解決できる。また、本発明によれば、適切な関連語の派生を行うことができないという問題を解決することができる。つまり、CPUの交換の例を用いて、本発明が、「手動または、共起関係を用いて作成した静的なシソーラスを使用したのでは適切な関連語の派生を行うことができな

い。」という問題を解決することが可能である。

【0060】その理由を以下に説明する。ここでは、キーワード派生ルールは、まだ、何も登録されていないものとする。あるユーザが1回目の検索条件として“CPU”と入力して検索を実行したが、“CPU”を含む文書があまりにたくさんありすぎて、どれが「CPUの交換」に関連するものかわからなかったとする。このとき、本発明では、手動または、共起関係を用いて作成した静的なシソーラスを使用した場合と同様に、適切な関連語の派生を行うことはできない。しかしながら、このユーザが考えた末、検索条件として“交換”を追加したならば、

(CPU、交換、関連度“大”)

がキーワード派生ルールとして蓄積される。従って、その後、同じ意図を持った(CPUの交換に関する文書が欲しい)ユーザが、同様に検索条件として、“CPU”と入力したならば、今度は、関連語として“交換”がユーザに提示され、検索実行処理(図3、ステップ150)で、“CPU⇔交換”の検索が行われる。このように本発明では、動的にシソーラスが構成されるので、手動または、共起関係を用いて作成した静的なシソーラスを使用した場合と比べて、適切な関連度の派生を行うことができる。

【0061】また、本発明では、検索条件の誤った交換規則が保存される問題を解決することができる。詳しくは、本発明のキーワード派生ルールは、ユーザが実際にこのルールを使用したか否かの情報を用いて動的に修正され、検索の結果には依存しない。そのため、例えば、意味的に正しくないキーワード派生ルールが保存されたとしても、その関連語はユーザによって選択されないの
30 で、関連度がしだいに小さくなり、その影響はしだいに小さくなる。従って、検索条件の誤った交換規則が保存されてしまう問題を解決できる。

【0062】また、上記の実施例は、図3、図4、及び図5に基づいて説明したが、この例に限定されことなく、図4、図5に示す構成及び図3に示す一連の動作をプログラムとして構築し、ユーザ側の装置または、情報検索装置として利用されるコンピュータに接続されるディスク装置、フロッピーディスク、CD-ROM等の可

搬記憶媒体に格納しておき、本発明を実施する際にインストールすることにより、容易に本発明を実現することができる。

【0063】なお、本発明は、上記の実施例に限定されことなく、特許請求の範囲内で種々変更・応用が可能である。

【0064】

【発明の効果】上述のように、本発明によれば、ユーザが入力したキーワードが適切なものでない場合でも、適切な検索を実行することができる。また、ユーザの検索を支援する機能を適切に働かせることが可能となる。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明の検索の概要を説明するためのフローチャートである。

【図4】本発明の一実施例の情報検索装置の構成図である。

【図5】本発明の一実施例の情報検索装置のユーザインタフェースの画面の例である。

【符号の説明】

200 ユーザ検索条件入力手段

210 形態素解析部

220 WWW処理部、ユーザ選択指示手段

225 検索結果表示手段

230 検索条件生成部

240 情報検索制御部

250 検索処理部/結果変換部、検索手段

255 再検索手段

30 260 派生ルール修正処理部、ルールベース修正手段

270 キーワード派生処理部、他キーワード取得手段

280 検索対象データベース

290 ルールベース

310 追加検索条件入力部

320 前回検索条件表示部

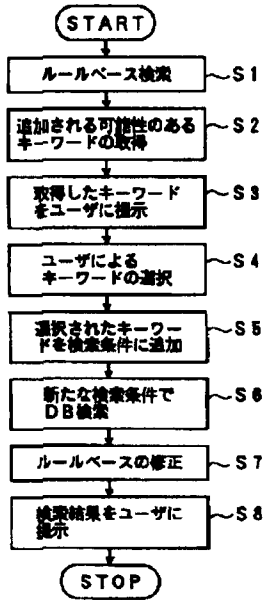
330 派生キーワード表示・選択部

340 検索種別選択部

350 検索結果表示部

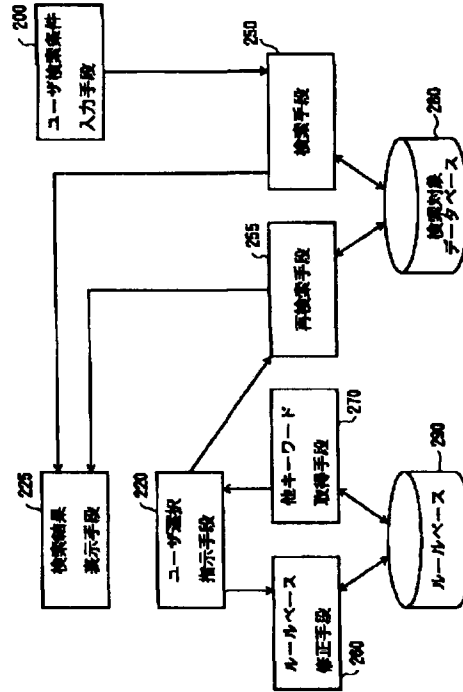
【図1】

本発明の原理を説明するための図



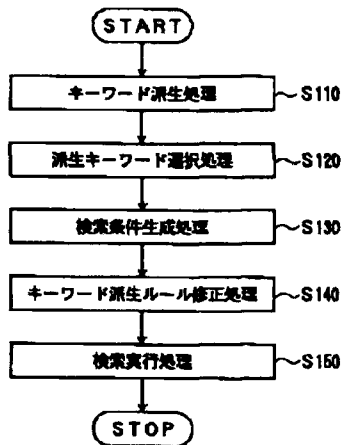
【図2】

本発明の原理構成図



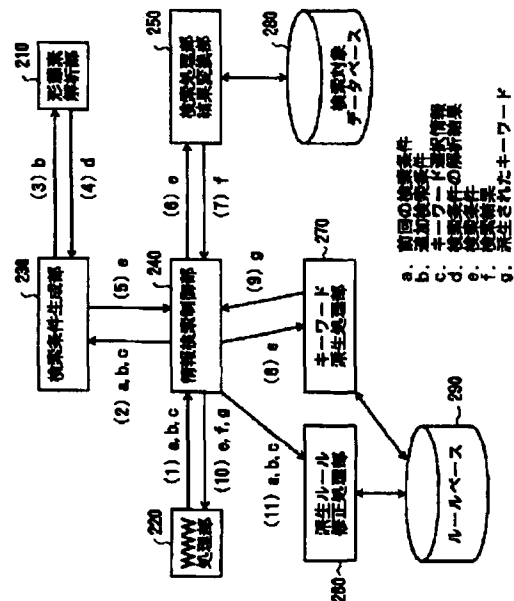
【図3】

本発明の検索の概要を説明するためのフローチャート



【図4】

本発明の一実施例の情報検索装置の構成図



【図5】

本発明の一実施例の情報検索装置の
ユーザインタフェースの画面の例

検索条件を追加してください

310 追加検索条件
入力部

検索実行

320 前回検索条件
表示部

入力語句：運転免許
検索条件：(運転 OR 免許)

330 派生キーワード
表示・選択部

派生キーワード
☐取得 ☐更新 ☐自動車 ☐手続書
☐交付 ☐免許証 ☐役所

340 検索種別選択部

検索オプション
☒絞り込み検索
☐再検索

350 検索結果表示部

検索結果は 47 件です
 1. 身体障害者に対する自動車運転免許証取得
 2. 運転免許の更新方法